इंटरनेट मानक

Disclosure to Promote the Right To Information

Whereas the Parliament of India has set out to provide a practical regime of right to information for citizens to secure access to information under the control of public authorities, in order to promote transparency and accountability in the working of every public authority, and whereas the attached publication of the Bureau of Indian Standards is of particular interest to the public, particularly disadvantaged communities and those engaged in the pursuit of education and knowledge, the attached public safety standard is made available to promote the timely dissemination of this information in an accurate manner to the public.

"जानने का अधिकार, जीने का अधिकार"
Mazdoor Kisan Shakti Sangathan
"The Right to Information, The Right to Live"

"पुराने को छोड नये के तरफ"
Jawaharlal Nehru
"Step Out From the Old to the New"

IS 14720 (1999): Documentation - Guidelines for the Establishment and Development of Monolingual Thesauri [MSD 5: Documentation and Information]

"ज्ञान से एक नये भारत का निर्माण"
Satyanarayan Gangaram Pitroda
"Invent a New India Using Knowledge"

Act Number 22 of 2005
The Indian Parliament
RIGHT TO INFORMATION

Made Available By
Public.Resource.Org

"ज्ञान एक ऐसा खजाना है जो कभी चुराया नहीं जा सकता है"
Bhartṛhari—Nītiśatakam
"Knowledge is such a treasure which cannot be stolen"

BLANK PAGE

BRUTUM FULMEN

IGNORANTIA JURIS NON EXCUSAT

PROTECTED BY COPYRIGHT

IS 14720 : 1999
ISO 2788 : 1986

(Reaffirmed 2004)

भारतीय मानक

प्रलेखन — एकभाषी विश्वकोश के निर्धारण और विकास के मार्गदर्शी सिद्धांत

Indian Standard

DOCUMENTATION — GUIDELINES FOR THE ESTABLISHMENT AND DEVELOPMENT OF MONOLINGUAL THESAURI

ICS 01.140.20

© BIS 1999

BUREAU OF INDIAN STANDARDS
MANAK BHAVAN, 9 BAHADUR SHAH ZAFAR MARG
NEW DELHI 110002

July 1999

Price Group 11

Documentation and Information Sectional Committee, MSD 5

NATIONAL FOREWORD

This Indian Standard which is identical with ISO 2788 : 1986 'Documentation – Guidelines for the establishment and development of monolingual thesauri' issued by the International Organization for Standardization (ISO) was adopted by the Bureau of Indian Standards on the recommendation of the Documentation and Information Sectional Committee (MSD 5) and approval of the Management and Systems Division Council.

The text of the ISO Standard has been approved as suitable for publication as Indian Standard without deviations. Certain conventions are, however, not identical to those used in Indian Standards. Attention is particularly drawn to the following:

a) Wherever the words 'International Standard' appear referring to this standard, they should be read as 'Indian Standard'.

b) Comma (,) has been used as a decimal marker while in Indian Standards, the current practice is to use a point (.) as the decimal marker.

In this adopted standard, references have also been made to the following International Standards for which there are no Indian Standards :

ISO 5963	Documentation – Methods for examining documents, determining their subjects and selecting indexing terms
ISO 5964	Documentation – Guidelines for the establishment and development of multilingual thesauri

The technical committee responsible for the preparation of this standard has reviewed the provisions of the above referred ISO standards and has decided that they are acceptable for use in conjunction with this standard.

At the time of publication, the editions of standards mentioned above, were valid. All standards are subject to revision, and users of these standards are advised to investigate the possibility of applying the most recent editions of the standards.

Indian Standard

DOCUMENTATION — GUIDELINES FOR THE ESTABLISHMENT AND DEVELOPMENT OF MONOLINGUAL THESAURI

0 Introduction

The effectiveness of a subject index as a means for identifying and retrieving documents depends upon a well-constructed indexing language. This applies to any system where the selection of indexing terms calls for human intellectual decisions, including those systems in which a computer is used to store and manipulate terms or to identify documents associated with terms or combinations of terms assigned by an indexer.

The compiler of a subject index faces three main tasks

a) determining the subject matter of documents;

b) selecting the terms which together summarize the subject;

c) indicating relationships between the concepts represented by these terms.1)

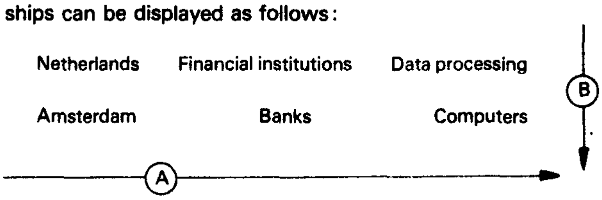
The first of these tasks is described separately in ISO 5963. The second and third tasks concern not only the indexer but also the user of the index. This International Standard deals with some aspects of term selection, since it contains recommended procedures for vocabulary control, but it is particularly concerned with means for establishing and displaying certain kinds of relationships between indexing terms.

Two kinds of inter-term relationships can be distinguished:

a) syntactical or *a posteriori* relationships between the terms which together summarize the subject of a document. For example, an indexer dealing with a work on "Computers in banks in Amsterdam" may assign three terms, "Banks", "Computers" and "Amsterdam", to the document. In a post-coordinate system the relationship between these terms is not explicitly indicated, and the document would be retrieved if any or all of these terms were used as retrieval keys. In a pre-coordinated index the relationships between the terms may be conveyed in various ways, for example by symbols which express specific relationships, the positions of terms within entries, their typography and/or their accompanying punctuation. The terms in this example are not normally associated according to common frames of reference, and their interrelationships can therefore be regarded as document-dependent;

b) those *a priori* or thesaural relationships between terms assigned to documents and other terms which, because they form part of common and shared frames of reference, are present by implication. In the example above, "Banks" would imply a broader term such as "Financial institutions"; "Computers" is mentally associated with "Data processing"; and "Amsterdam" implies the wider location "Netherlands". Any of these mentally-associated terms might serve as a user's approach to the subject index. These relationships are document-independent, since they are generally recognized and could be established through reference to standard works, such as dictionaries and encyclopaedias.

The distinction between these two kinds of inter-term relationships can be displayed as follows:



Ⓐ = *a posteriori* relationships between indexing terms *assigned to a document*

Ⓑ = *a priori* relationships handled by the thesaurus

This International Standard is especially concerned with those *a priori* relationships which can be displayed in a thesaurus, where they then, in effect, add a second dimension to an indexing language, as shown above.

1 Scope and field of application

1.1 The recommendations set out in this International Standard are intended to ensure consistent practice within a single indexing agency, or between different agencies (for example members of a network). They should not be regarded, however, as mandatory instructions. Optional procedures are described in many cases, for example for the display of inter-term relationships, without indicating one of these approaches as the preferred technique. The choice of procedure will vary from one indexing agency to another, depending on management decisions that fall outside the scope of this International

---

1) For practical purposes, "term" and "concept" are sometimes used interchangeably.

Standard. As far as possible the techniques described in this International Standard are based upon general principles which apply to any subject field. It is recognized, however, that an indexer working within a limited subject field may sometimes need to depart from these general recommendations, and this is noted where appropriate.

1.2 As far as possible the techniques described in this International Standard are not limited to a particular method of indexing, whether post-coordinate or pre-coordinate. This International Standard is, however, subject to the following restrictions:

a) it deals with the display and organization of terms that form a controlled subset of natural language. It does not suggest procedures for organizing and displaying mathematical and chemical formulae;

b) it is generally based on the concept of "preferred terms" (see 3.5);

c) its application is limited to agencies in which human indexers are used to analyse documents and express their subjects in the terms of a controlled indexing language. It is not applicable to these agencies which apply entirely automatic indexing techniques, where terms occurring in texts are organized into sets according to criteria which can be established by a computer, for example their frequency of occurrence and/or adjacency in the text. It is considered, however, that a well-constructed monolingual thesaurus can serve as a useful aid when searching such a free-text system;

d) it deals mainly with procedures for indexing collections of documents listed in catalogues or bibliographies. It is not concerned with the preparation of "back-of-the-book" indexes, although many of its recommended procedures may be useful for that purpose.

1.3 The recommendations contained in this International Standard relate to monolingual thesauri, without reference to the special requirements of multilingual thesauri, i.e. those thesauri in which conceptual equivalences are expressed in terms selected from more than one natural language. The construction and maintenance of a multilingual thesaurus is dealt with separately in ISO 5964. Since the principles on which this International Standard is based can be regarded as both language-independent and culture-independent, they have also been accepted as the basis for the multilingual standard. Consequently, general principles and procedures which apply equally to both kinds of thesauri are fully described only in this International Standard, i.e. they are not repeated in ISO 5964.

2 References

ISO 5963, *Documentation — Methods for examining documents, determining their subjects and selecting indexing terms.*

ISO 5964, *Documentation — Guidelines for the establishment and development of multilingual thesauri.*

3 Definitions

For the purposes of this International Standard, the following definitions apply.

3.1 document: Any item, printed or otherwise, that is amenable to cataloguing and indexing.

NOTE — This definition refers not only to written and printed materials in paper or microform versions (for example books, journals, diagrams, maps), but also to non-printed media (for example machine-readable records, films, sound recordings, etc.), and three-dimensional objects or realia used as specimens.

3.2 indexing language: A controlled set of terms selected from natural language and used to represent, in summary form, the subjects of documents.

3.3 thesaurus: The vocabulary of a controlled *indexing language* (see 3.2), formally organized so that the *a priori* relationships between concepts (for example as "broader" and "narrower") are made explicit.

3.4 indexing term: The representation of a concept, preferably in the form of a noun or noun phrase.

NOTE — An indexing term can consist of more than one word, and is then known as a *compound term* (see 3.7). In a controlled indexing vocabulary, a term is designated either as a *preferred term* or as a *non-preferred term.*

3.5 preferred term: A term used consistently when indexing to represent a given concept; sometimes known as "descriptor".

3.6 non-preferred term: The synonym or quasi-synonym of a preferred term. A non-preferred term is not assigned to documents, but is provided as an entry point in a thesaurus or alphabetical index, the user being directed by an instruction (for example USE or SEE) to the appropriate preferred term; sometimes known as "non-descriptor".

3.7 compound term: *An indexing term* (see 3.4) which can be factored morphologically into separate components, each of which could be expressed, or re-expressed, as a noun that is capable of serving independently as an indexing term.

NOTE — The parts of the great majority of compound terms can be distinguished as follows:

a) the **focus** or **head**, i.e. the noun component which identifies the general class of concepts to which the term as a whole refers.

Examples:

1) the noun component "indexes" in the compound term "printed indexes".

2) the noun "hospitals" in the prepositional phrase "hospitals for children".

b) The **difference** or **modifier**, i.e. one or more further components which serve to narrow the extension of the focus and so specify one of its subclasses.

Examples:

1) the adjective "printed" in the compound term "printed indexes".

2) the preposition-plus-noun combination "for children" in the compound term "hospitals for children".

The focus and its difference(s) may be written as separate words, as in "dining rooms" and "soup spoons", or they may be concatenated into single words, as in "bedrooms" and "teaspoons".

3.8 node label: A "dummy" term not assigned to documents when indexing, but inserted into the systematic section of some types of thesauri to indicate the logical basis on which a category has been divided; sometimes known as a "facet indicator".

Examples:

By occupation

By purpose

Parts

NOTE — See 8.3.3 for a further description of node labels.

4 Abbreviations and symbols

4.1 The following abbreviations, which are used throughout this International Standard, are printed as prefixes to terms, etc. Each abbreviation indicates the relationship or function of the term or note which follows, as explained below:

SN	Scope note; a note attached to a term to indicate its meaning within an indexing language
USE	The term that follows the symbol is the preferred term when a choice between synonyms or quasi-synonyms exists
UF	Use for; the term that follows the symbol is a non-preferred synonym or quasi-synonym
TT	Top term; the term that follows the symbol is the name of the broadest class to which the specific concept belongs; sometimes used in the alphabetical section of a thesaurus
BT	Broader term; the term that follows the symbol represents a concept having a wider meaning
BTG	Broader term (generic)
BTP	Broader term (partitive)
NT	Narrower term; the term that follows the symbol refers to a concept with a more specific meaning
NTG	Narrower term (generic)
NTP	Narrower term (partitive)
RT	Related term; the term that follows the symbol is associated, but is not a synonym, a quasi-synonym, a broader term or a narrower term

4.2 Abbreviations with equivalent meanings also occur in thesauri in other languages.

Examples:

French

NE	Note explicative
EM	Employer
EP	Employé pour
MV	Nom de la classe la plus générale
TG	Terme générique
TGG	Terme générique générique
TGP	Terme générique partitif
TS	Terme spécifique
TSG	Terme spécifique générique
TSP	Terme spécifique partitif
VA	Voir aussi

German

D	Definition
BS	Benutzen
BF	Benutzt für
SB	Spitzenbegriff
OB	Oberbegriff
OA	Oberbegriff (Abstraktionsrelation)
SP	Verbandsbegriff (Bestandsrelation)
UB	Unterbegriff
UA	Unterbegriff (Abstraktionsrelation)
TP	Teilbegriff (Bestandsrelation)
VB	Verwandter Begriff

4.3 The abbreviations listed in 4.1 and 4.2 have acquired status as generally recognized conventions, and they occur in many published thesauri. They have obvious mnemonic value, yet it is realized that they are also language-dependent. If this characteristic is regarded as sufficiently important to justify using a "neutral" system, an agency can adopt the language-independent symbols developed by ISO and given in the annex.

4.4 The following conventions are also used in examples throughout this International Standard:

a) preferred terms are printed in upper case throughout:

Examples:

CARS

ANIMALS

b) non-preferred terms are printed in lower case except when the non-preferred term is a proper name requiring an upper case initial, or an abbreviation or acronym which should be printed throughout in upper case:

Examples:

CARS
  UF automobiles

ANIMALS
  UF fauna

WORLD HEALTH ORGANIZATION
  UF WHO

4.5 A compound term is sometimes factored morphologically into components, and these are assigned independently as indexing terms (see clause 7). If the unfactored form is likely to be sought by users, a reference should be made from the compound term as a whole to the separate terms used in combination.

Example:

coal mining **USE** COAL + MINING

5 Vocabulary control

5.1 Two principal means for achieving vocabulary control are employed in thesauri:

a) terms are deliberately restricted in scope to selected meanings. Unlike the terms in a dictionary, which may be accompanied by a number of different definitions reflecting common usage, each term in a thesaurus is generally restricted to whichever single meaning serves the needs of an indexing system most effectively. The structure of a thesaurus, notably its display of hierarchical relationships, frequently indicates the intended meaning of a term. If this technique is not sufficiently explicit, a definition or scope note should be appended to the term. This note should state the chosen meaning, and it may also indicate other meanings which are recognized in natural language but which have been deliberately excluded for indexing purposes;

b) when the same concept can be expressed by two or more synonyms, one of these terms is usually selected as the preferred term (see 3.5) which is then used consistently in indexing. Reference to the preferred term should be made from any synonym which might also function as a user's access point.

5.2 Further means for achieving vocabulary control are reviewed in the following clauses. These deal with matters such as the choice of singular or plural forms, the selection of the preferred term when synonyms are encountered, and the extent to which a compound term should either be retained in its pre-coordinated form or factored into separate components, each expressed as a noun and used independently as an indexing term.

6 Indexing terms

6.1 General

6.1.1 The concepts represented by indexing terms belong to general categories such as the following:

a) **Concrete entities**

1) Things and their physical parts

Examples:

BIRDS

LIMBS

MICROFORMS

MOUNTAIN REGIONS

2) Materials

Examples:

ADHESIVES

RUBBER

TITANIUM

b) **Abstract entities**

1) Actions and events

Examples:

GLACIATION

GOLF

MARKETING

2) Abstract entities, and properties of things, materials or actions

Examples:

ELASTICITY

NEWS

PERSONALITY

SPEED

3) Disciplines or sciences

Examples:

ARCHAEOLOGY

CHEMISTRY

4) Units of measurement

Examples:

HERTZ

KILOMETRES

c) **Individual entities, or "classes-of-one", expressed as proper nouns**

Examples:

SRI LANKA

WORLD HEALTH ORGANIZATION

6.1.2 The compiler of a thesaurus needs to be aware of these classes, since they are likely to affect some of the procedures considered in later clauses, for example the choice of singular or plural forms, and applying a test for the validity of a hierarchy.

6.2 Forms of terms

6.2.1 Nouns and noun phrases

An indexing term should preferably consist of a noun or a noun phrase. Noun phrases belong to the category of compound terms, and occur in two forms:

a) **adjectival phrases**

Example:

MARINE BIRDS

This class also includes those single-word compounds which can be factored morphologically into a noun plus a modifying difference having an adjectival function:

Examples:

FOOTBALL

MOTORWAYS

b) **prepositional phrases**

Example:

HOSPITALS FOR CHILDREN

Those parts of a compound term which function as differences [see item b) of the note to 3.7] should be considered as potential sources of extra terms in a thesaurus. When a difference consists of an adjective, the noun from which the adjective was derived should be preferred as the extra candidate term. If these terms are accepted, the thesaurus should display reciprocal relationships between the extra term and the compound term as a whole.

Examples:

a) MARINE BIRDS
RT SEAS

SEAS
RT MARINE BIRDS

b) SCHOOLS FOR HANDICAPPED CHILDREN
RT HANDICAPPED CHILDREN

HANDICAPPED CHILDREN
RT SCHOOLS FOR HANDICAPPED CHILDREN

6.2.2 Adjectives

6.2.2.1 Adjectives used alone may occur in an indexing language in special circumstances considered below, but their use should be avoided as far as possible.

6.2.2.2 Adjectives may be accepted as single words in an index or thesaurus in situations such as the following:

a) when working in a language where adjectives generally precede the nouns they modify, the user may be directed, on economic grounds, from a noun to an adjective which serves as the first component of several compound terms. For example, a reference could be made from "France" (the noun) to "French" (the adjective) if the indexing language contains a number of terms such as "French art", "French language", "French literature", "French wines". This would apply especially when the adjective and the noun from which it was derived differ widely in their spellings, for example France/French, Sea/Marine;

b) in languages where adjectives follow the nouns they determine, a reference may be made from an adjective to one or more noun phrases containing the adjective. An index in French, for example, could contain reference from an adjective such as "Pasteurisé" to compound terms such as "Crème pasteurisée", "Lait pasteurisé", and "Produits pasteurisés".

6.2.3 Adverbs

Adverbs such as "Very" or "Highly" should not be used alone as indexing terms. A phrase beginning with an adverb should not be accepted as an indexing term except when it has acquired a special meaning within a jargon.

Example:

VERY HIGH FREQUENCY

6.2.4 Verbs

Verbs expressed as infinitives or participles should not be used alone as indexing terms. Activities should be represented by nouns or verbal nouns.

Examples:

COOKERY (not "cook")

DISTILLATION (not "distil")

6.2.5 Abbreviations and acronyms

Abbreviations and acronyms should not be used as preferred terms except when they are widely used and readily understood within the field covered by the thesaurus. Many acronyms and abbreviations can refer to more than one concept, and the full form of the name should therefore function as the preferred term, with a reciprocal reference from the abbreviated form.

Example:

WORLD HEALTH ORGANIZATION
  UF WHO

WHO **USE** WORLD HEALTH ORGANIZATION

Abbreviations and acronyms may function as preferred terms if they have become so well established that the full form of the name is rarely used or is generally ignored. Reciprocal references should still be made between the full term and its abbreviation.

Example:

UNICEF
  UF United Nations International Children's Fund

United Nations International Children's Fund
  USE UNICEF

6.3 Choice of singular or plural forms

6.3.1 In those languages where a distinction between singulars and plurals can be made, the decision to adopt singular or plural forms as indexing terms is likely to be affected by factors such as the following:

a) post-coordinate or pre-coordinate indexing

In a pre-coordinate index, terms selected from a thesaurus are organized into index entries in such a way that the entry as a whole expresses a subject in summary form. Relationships between terms may be conveyed in various ways, for example by word order and/or the choice of special typography and punctuation. In some systems, terms may be organized into phrases linked by prepositions or other adjuncts. In these circumstances, the meaning or comprehensibility of the index entry as a whole may be affected be the use of singular or plural forms. This does not apply to a post-coordinate system, where terms are assigned to documents as independent retrieval keys, without indicating their interrelationships.

b) cultural factors

Agencies within different countries tend to work within different traditions concerning the use of singulars or plurals. In English-speaking countries, for example, terms may be expressed either as singulars or plurals depending upon factors considered below (see 6.3.2). Indexers in other language communities, for example French and German, tend to prefer the singular form where possible, so that the user can approach the thesaurus or index in the same way as a dictionary. In these cases, however, exceptions to the general preference are likely to occur for pragmatic reasons depending, for example, upon the type of indexing system used [see a) above], or an occasional need to avoid ambiguity where the singular form can refer to more than one concept, and these could be distinguished by expressing one of them as a plural.

6.3.2 In agencies where either the singular or plural form of a term may be adopted, the choice of the preferred term is generally related to the kind of concept to which the term refers. As noted earlier (see 6.1.1) terms can be divided into those that represent concrete entities, and those that refer to abstract concepts. These two classes are reviewed separately below.

6.3.2.1 Nouns that represent concrete entities can be divided into two further categories:

a) count nouns, i.e. names of countable objects that are subject to the question "How many?" but not "How much?". These should be expressed as plurals.

Examples:

DOCUMENTS

PENGUINS

POLITICAL PARTIES

WINDOWS

Special treatment is usually given to the names of parts of the body. These should be expressed as plurals when more than one occurs in a fully formed organism, but in the singular if only one is present.

Examples:

EARS	*but*	DIGESTIVE SYSTEM
HANDS		HEAD
LUNGS		NOSE

b) non-count nouns, for example names of materials or substances which are subject to the question "How much?" but not "How many?". These should be expressed as singulars.

Examples:

PAINT

QUARTZ

STEAM

If the community of users served by the index regards a given substance or material as a class with more than one member, the class should be expressed in the plural.

Examples:

PLASTICS

POISONS

6.3.2.2 The names of abstract concepts, for example abstract entities and phenomena, properties, systems of belief, activities and disciplines, should be expressed in their singular forms:

Examples:

Abstract entities and phenomena: PERSONALITY; WINTER

Properties: BRITTLENESS; OPACITY; SOLUBILITY

Systems of belief: CATHOLICISM; SHINTOISM

Activities: CUTTING; IMMIGRATION; RESPIRATION

Disciplines: PHYSICS; SOCIOLOGY

When an abstract concept is regarded as a class with more than one member, the term representing the class should be expressed in the plural.

Examples:

CHEMICAL REACTIONS

PHYSICAL SCIENCES

SETS

6.3.3 Where the singular and plural forms of a term refer to different concepts both should be entered in the thesaurus. If necessary, the distinction should be indicated by a qualifying term or phrase.

Examples:

WOOD (material)

WOODS (areas of woodland)

Note that the added qualifier then becomes an integral part of the term; it does not constitute a scope note (see also 6.6).

6.3.4 Where the spelling of the singular and plural differs to such an extent that the terms would be separated by unrelated terms when filed alphabetically, a reference should be made from the non-preferred form.

Example:

mouse **USE** MICE

6.4 Homographs or polysemes

Homographs or polysemes (sometimes referred to by the broader term "homonyms") are words with the same spelling but different meanings.

Example:

CRANES, which can refer to either birds or lifting equipment.

When homographs are encountered in indexing, each should be supplemented by a qualifying word or phrase. The indexing term should be distinguished from its qualifier, for example by using a different typeface, or by inserting the qualifier between parentheses. The qualifier does not serve as a scope note (see 6.6), and the term and its qualifier should be assigned to the thesaurus as a unit.

Examples:

BEAMS (radiation)

BEAMS (structures)

CRANES (birds)

CRANES (lifting equipment)

6.5 Choice of terms

6.5.1 Spelling

The most widely accepted spelling of words should be adopted. If variant spellings exist and are commonly recognized, each should be entered in the thesaurus, and a reference should be made from the non-preferred to the preferred form.

Example:

Rumania **USE** ROMANIA

Roumania **USE** ROMANIA

Where possible, spelling should follow the practice of a well-established dictionary or glossary. If a choice between spellings is made for cultural reasons (for example between American English and British English), the chosen source should be stated in the Introduction, and the choice should be adhered to consistently throughout the thesaurus.

6.5.2 Loan words and translations of loan words

Terms from other languages are sometimes encountered as "loan words", i.e. foreign terms which are accepted as newly-coined terms. If these terms are well-established, they should be incorporated into the thesaurus. Occasionally a loan word and a putative translation co-exist. If the loan word is more widely accepted, it should be treated as the preferred term, but if the translation becomes well-established, this should be preferred. Reciprocal references should be made between the preferred and non-preferred terms.

Example:

X-RAYS
  UF roentgen rays

roentgen rays **USE** X-RAYS

6.5.3 Transliteration

When transliterating terms from languages with different alphabets, the procedure described in a relevant International Standard should be followed as far as possible.

6.5.4 Slang terms and jargon

Slang and jargon may be encountered when indexing in circumstances such as the following:

a) A newly-emerging concept is expressed by a term originating within a particular subculture or social group, and no widely accepted alternative exists. The slang or jargon term should then be accepted as an indexing term.

Example:

BYTE

HIPPIES

b) A slang or jargon term emerges as an alternative to an existing and well-established term. The established term should then be chosen as the preferred term, and the slang term should be admitted as a non-preferred term only if it is so widely recognized that it might serve as the user's access point to the index.

Example:

ASSOCIATION FOOTBALL
  UF soccer

soccer **USE** ASSOCIATION FOOTBALL

6.5.5 Common names and trade names

A product is frequently known by a widely recognized trade name. Where a suitable common name also exists, this should be adopted as the preferred term, and the trade name should be admitted as a non-preferred term only if it is likely to serve as a user's access point to the index.

Example:

POLYETHYLENE
  UF polythene

polythene **USE** POLYETHYLENE

6.5.6 Popular names and scientific names

If a popular and a scientific name refer to the same concept, the form most likely to be sought by the users of the index should be chosen. For example, "Penguins" might be chosen as the preferred term in a general index, but the scientific equivalent, "Sphenisciformes", may be preferred in a zoological index. Reciprocal references should be made in these cases.

6.5.7 Place names

Names of countries and geographical regions sometimes vary from language to language. Variant terms referring to the same place also occur within a single language community, for reasons such as the following:

a) an "official" and a "popular" name are both in common use:

Example:

HOLLAND

NETHERLANDS

b) The original and a vernacular form co-exist:

Example:

LEGHORN

LIVORNO

The name which is most familiar to the users of the thesaurus should be designated as the preferred term. Other things being equal, preference should be given to the official rather than the popular name. Reciprocal references should be made between the preferred and non-preferred versions.

Example:

NETHERLANDS
  UF Holland

Holland **USE** NETHERLANDS

6.5.8 Proper names of institutions and persons

These names are frequently excluded from a thesaurus. When they are included, the form of name should be selected in accordance with a recognized code of cataloguing practice. These names can sometimes vary, depending upon their origins, in which case the following should be followed:

a) names of national or local institutions, i.e. those that conduct their business, and perhaps publish journals or proceedings, in only one language, should be recorded in their untranslated forms. If a translated form also exists, this should be designated as a non-preferred term.

Example:

ACADEMIE DES SCIENCES
  UF Academy of Sciences (Paris)

Academy of Sciences (Paris)
USE ACADEMIE DES SCIENCES

b) names of international organizations, or of local institutions which publish documents in more than one language, should be expressed in their best-known forms. If the index is used by readers from different language communities who might not know the "local" name of an international organization, vernacular forms from other languages may be admitted as non-preferred terms.

Example (in an English-language thesaurus):

UNITED NATIONS
UF Nations unies
UN

Nations unies **USE** UNITED NATIONS

UN **USE** UNITED NATIONS

c) personal names should usually be recorded in their original forms. Some individuals, notably historical figures who have acquired international recognition, have local forms of name which are so well-established that they should be accepted as preferred terms, references being made where necessary from the original names. It should be realized, however, that this is an exceptional case.

Example (in an English-language thesaurus):

JOAN OF ARC
UF Jeanne d'Arc

Jeanne d'Arc **USE** JOAN OF ARC

6.6 Scope notes and definitions

6.6.1 Scope notes and definitions may be appended to a term to indicate the limited sense in which it is used for indexing purposes and thus exclude other possible meanings. Further kinds of information may also be indicated in a scope note, for example:

a) the date when a term was adopted, or when its scope was redefined;

b) the source of the term, especially in the case of neologisms, i.e. newly-coined terms which represent emergent concepts;

c) instructions intended for indexers, concerning, for example, combinations of terms which are allowed or not allowed within a given indexing language.

Scope notes should be distinguished typographically from the terms to which they are appended.

Example:

MICROWAVE FREQUENCIES
SN 1 GHz to 300 GHz

6.6.2 Unlike the qualifiers considered in 6.3.3 and 6.4, a scope note is not regarded as forming part of the term to which it is attached. It is not intended to be a dictionary definition, but it serves instead to indicate the use of a term within a given indexing language. It is not necessary to append scope notes or definitions to all the terms in a thesaurus. A thesaurus displays, through its structure, relationships which generally establish the "meaning" context of a given term, with particular reference to terms of wider and narrower connotation. This is usually sufficient to indicate the intended interpretation of a particular term. A scope note occasionally needs to be expanded into a full definition, if, for example, a term is interpreted loosely in common usage, or if different dictionaries give variant meanings.

Example:

DOCUMENTATION
SN the process of storing and retrieving information in all fields of knowledge

7 Compound terms

7.1 General

7.1.1 The establishment of procedures for dealing consistently with compound terms introduces one of the most difficult areas in the field of subject indexing. It can be stated, as a general rule, that terms should represent simple or unitary concepts as far as possible, and compound terms should be factored into simpler elements except when this is likely to affect the user's understanding. Complex subjects should be expressed by combinations of separate terms, and these may be assigned as independent search keys in a post-coordinate system, or they may function as components of pre-coordinated index entries. It is realized, however, that this general recommendation does not stipulate the exact circumstances in which a compound term encountered in documents should be factored into separate components, or else retained in its compound form, for indexing purposes. The procedures described below are offered as guidelines for indexers using controlled vocabularies who need criteria to ensure consistent practice in the treatment of compound terms. Within this context, it is considered that the establishment of such criteria constitutes an important step towards achieving consistency, and hence predictability, not only within a single indexing agency, but also among different agencies working together, for example as members of a network or as contributors to a common data base.

7.1.2 Compound concepts can frequently be expressed in more than one form. For example:

a) in English, the concept represented by "children's hospitals" can also be represented by the prepositional phrase "hospitals for children";

b) in French, the concept represented by "pollution de l'air" can also be expressed as "pollution atmosphérique".

The procedures reviewed in this International Standard are not concerned with the choice between such forms when more than one exists. The procedures described below refer to criteria for determining whether a given compound term should be retained in its pre-coordinated form or factored into separate components, each of which is accepted as an indexing term in its own right.

7.1.3 Two techniques for factoring (i.e. analysing a term into separate meaning elements) are recognized in indexing:

a) semantic factoring

This technique can be applied to single words as well as to compound terms. A term which represents a complex notion is re-expressed in the form of simpler or definitional elements, each of which can also occur in other combinations to represent a range of different concepts.

Example:

thermometers could be expressed by a combination of three terms:

TEMPERATURE + MEASUREMENT + INSTRUMENTS

This technique is not recommended. It can be employed only in post-coordinate indexing systems, and it should be applied only to terms in fringe subject areas by agencies which need to limit the size of their indexing vocabularies. It is generally recognized that semantic factoring leads to a loss of precision in retrieval.

b) syntactical factoring

This technique is applied to compound terms, i.e. terms which are amenable to morphological analysis into separate components, each of which can be accepted as an indexing term in its own right.

Example:

building construction expressed as

BUILDINGS + CONSTRUCTION

If syntactical factoring is applied to a term in common use which may serve as a user's access point to the index, the term as a whole should be entered in the thesaurus as a non-preferred term, and a reference should be made to the elements used in combination.

Example:

building construction
USE BUILDINGS + CONSTRUCTION

7.1.4 In explaining the grounds for deciding whether a term should be retained in its pre-coordinated form or syntactically factored, it is necessary to distinguish between the parts of a compound term through reference to their roles and interrelationships. The majority of compound terms, whether adjectival phrases or prepositional phrases, can be analysed into the following components:

a) the focus (also known in linguistics as the "head"): the noun component which identifies the broader class of things or events to which the term as a whole refers.

Examples:

"CONCRETE" in the adjectival phrase
"REINFORCED CONCRETE"

"HOSPITALS" in the prepositional phrase
"HOSPITALS FOR CHILDREN"

b) the difference (also known as the modifier): that part of a compound term which refers to a characteristic, or logical difference, which when applied to the focus narrows its connotation and so specifies one of its subclasses.

Examples:

"REINFORCED", which specifies a subclass of concrete in "REINFORCED CONCRETE".

"FOR CHILDREN", which specifies a subclass of hospitals in "HOSPITALS FOR CHILDREN".

7.1.5 A small proportion of terms encountered in indexing are not amenable to this kind of logical analysis. These are terms in which one of the elements appears to function as a difference, but it does not, as in the examples above, specify a subclass of the focal concept; instead, it tends to contradict membership of that class, to the extent that "artificial flowers" are not species of a class of things called "flowers", and "chocolate eggs" are not "eggs". The noun in these cases is described as syncategorematic, since it cannot stand alone as an indicator of the class of concepts to which the whole of the term refers. The term as a whole should then be treated as a single semantic unit.

7.1.6 These factors can be taken into account when deciding whether a concept should be expressed as a single compound term, or factored into two or more separate terms. Two approaches to decisions of this kind are reviewed below. Criteria which determine when a term should be retained in its compound form are surveyed in 7.2. Situations in which syntactical factoring is recommended are considered in 7.3.

7.2 Terms that should be retained as compounds

7.2.1 Compound terms should be retained as indexing terms in the following circumstances:

a) The compound term has become so familiar in common use, or in the field covered by the thesaurus, that its expression as separate elements would hinder comprehension:

Examples:

DATA PROCESSING

FINANCIAL MANAGEMENT

b) Syntactical factoring would lead to a loss of meaning, or ambiguity.

Example:

FOOD + PLANTS could represent either FOOD **for** PLANTS, or PLANTS **as** FOOD

7.2.2 More specifically, the following classes of term should be retained in their compound forms:

a) proper names, including terms that incorporate proper names:

Examples:

FREUDIAN ERRORS

UNITED NATIONS

b) terms in which the difference has lost its original meaning:

Examples:

LAWN TENNIS

ROUND TABLE CONFERENCE

TRADE WINDS

c) terms containing a difference which suggests a resemblance, as a simile, to an unrelated thing or event:

Examples:

CASCADE GENERATORS

TREE STRUCTURES

WING NUTS

d) terms which cannot be re-expressed (for example defined) without the use of an extra noun that is present in the compound only by implication:

Examples:

FIRE ESCAPES, i.e. ESCAPES for **PEOPLE** from FIRE

PETROL ENGINES, i.e. ENGINES using PETROL as a **FUEL**

e) terms containing syncategorematic nouns:

Examples:

ARTIFICIAL LIMBS

FOSSIL FISHES

7.3 Terms that should be syntactically factored

7.3.1 It is more difficult to specify exact criteria for recognizing those kinds of compound terms which should be factored into separate components, each of which is then entered as a noun or noun phrase in the thesaurus. The recommendations considered below are based upon general criteria, but they are not regarded as mandatory instructions to be applied rigidly in all circumstances. Some of the situations which call for terms to be retained in their unfactored forms were reviewed in 7.2.1 above. The following factors may also affect these decisions:

a) indexing in a special field, which may call for special treatment of terms which are familiar to users within the field covered by the thesaurus;

b) control over the number of terms assigned to documents, or the size of the indexing vocabulary.

In a post-coordinate system, the indexer may need to choose between the following procedures:

1) using compound terms to reduce the number of terms assigned (or "posted") to each document, but at the cost of a larger indexing vocabulary;

2) reducing the size of the indexing vocabulary by factoring terms into simpler elements, then assigning a higher number of terms to each document.

7.3.2 The following recommendations, which have been tested in various languages, are offered as a means for achieving consistent treatment of those compound terms which should be syntactically factored. Each of these recommendations is accompanied by a converse condition which can also be used to identify classes of terms that cannot usually be factored without a risk of ambiguity.

7.3.2.1

a) A compound term should be factored if the focus refers to a property or part (including materials), and the difference represents the whole or possessor of that property or part:

Examples:

In the terms

1) AIRCRAFT ENGINES

2) HOSPITAL PERSONNEL

3) SOIL ACIDITY

the focal nouns, i.e.

1) ENGINES

2) PERSONNEL

3) ACIDITY

represent parts or properties, and the differences

1) AIRCRAFT

2) HOSPITAL

3) SOIL

represent their possessing wholes.

These terms should therefore be re-expressed as separate nouns, i.e.

1) AIRCRAFT + ENGINES

2) HOSPITALS + PERSONNEL

3) SOILS + ACIDITY

b) Conversely, the name of the whole may be modified by the name of its part or property

Examples:

ACID SOILS

CONCRETE BRIDGES

SKILLED PERSONNEL

7.3.2.2

a) The name of a transitive action should not be modified by the name of the object on which the action was performed

Examples:

In the terms

1) OFFICE MANAGEMENT

2) STEEL CASTING

3) TEXTILE PRINTING

the focal nouns

1) MANAGEMENT

2) CASTING

3) PRINTING

represent transitive actions, and the differences

1) OFFICE

2) STEEL

3) TEXTILE

represent the patients of these actions.

These terms should therefore be factored, and the parts re-expressed as separate nouns:

1) OFFICES + MANAGEMENT

2) STEEL + CASTING

3) TEXTILES + PRINTING

b) Conversely, the name of a thing or material may be modified by the name of an action carried out upon it

Examples:

CAST STEEL

EXTRUDED PLASTICS

PRINTED TEXTILES

7.3.2.3

a) The name of an intransitive action should not be modified by the name of the performer of the action

Examples:

In the terms

1) BIRD MIGRATION

2) PLANT GROWTH

3) METAL DETERIORATION

the focal nouns

1) MIGRATION

2) GROWTH

3) DETERIORATION

represent intransitive actions, and the differences

1) BIRD

2) PLANT

3) METAL

represent the performers of these actions.

These terms should therefore be re-expressed as separate nouns:

1) BIRDS + MIGRATION

2) PLANTS + GROWTH

3) METAL + DETERIORATION

b) Conversely, the name of a thing may be modified by the name of the intransitive action in which it is or was engaged

Examples:

DEGRADED PLASTICS

MIGRATING BIRDS

MOVING PAVEMENTS

7.3.3 In addition to their role in achieving inter-indexer and inter-agency consistency, the recommendations above have been established with the following purposes in mind:

a) to avoid over-complexity in indexing terms;

b) to achieve a structured thesaurus based upon principles that are amenable to logical explanation and teaching.

In practice, these recommendations would ensure that a compound term such as "Aircraft engines" is factored, according to the procedure noted in 7.3.2.1, into two separate noun components, i.e. "aircraft", which belongs to a category of "vehicles", and "engines", which should be assigned to a different category, "prime movers". They also offer a basis for a standard treatment of compound terms where a given focus may be modified by more than one difference. If, for example, an indexing agency worked with policies which allow a phrase such as "aircraft engines" in addition to the phrases "jet engines" and "military aircraft", the indexer would face difficulties when dealing with a complex subject such as "military aircraft jet engines". In this kind of situation, the rules above offer a means of control by ensuring that the compound term is factored into separate and simpler components, i.e. "jet engines" and "military aircraft", each of which would be assigned to its own category in the thesaurus.

7.4 Order of words in compound terms

Both adjectival and prepositional noun phrases should preferably be entered in the thesaurus in natural language order, not as inverted terms.

Examples:

BIRDS OF PREY **not** PREY, BIRDS OF

ELECTRONIC CALCULATORS **not** CALCULATORS, ELECTRONIC

POLYSTYRENE CONTAINERS **not** CONTAINERS, POLYSTYRENE

When dealing with adjectival phrases in languages where nouns usually follow their adjectives, the focal nouns (for example "containers" and "calculators"), should be entered in their own right in the thesaurus, and they will then function as superordinate terms in hierarchies. The inverted form of a prepositional phrase may be entered as a non-preferred term, and a reference should then be made to the preferred term in its natural language order.

Examples:

matter, states of **USE** STATES OF MATTER

prey, birds of **USE** BIRDS OF PREY

8 Basic relationships in a thesaurus

8.1 General

8.1.1 A thesaurus should display and clearly distinguish between the basic relationships linking the terms it contains. Means for displaying relationships are described in clause 9; the present clause is concerned with defining and illustrating the relationships themselves.

8.1.2 Relationships of two different kinds can occur in a thesaurus:

a) relationships between individual concepts are reviewed in 8.2 to 8.4, where it will be seen that the hierarchical relationship effectively organizes terms into categories, subcategories, etc. according to their meanings;

b) in thesauri which contain systematic displays (see 9.3), the compiler needs to consider not only these inter-term relationships, but also relationships between categories, subcategories, etc. These inter-category relationships are considered separately in 9.3.2.

8.1.3 Three classes of basic inter-term relationship are recognized in thesauri:

a) the equivalence relationship (see 8.2);

b) the hierarchical relationship (see 8.3);

c) the associative relationship (see 8.4);.

Further subdivisions of each of these are described below. Each of these relationships is regarded as reciprocal, and this should be indicated by any system of symbols or abbreviations used to represent relationships in a thesaurus. The conventional abbreviations described in clause 4 are used in the examples that follow. An alternative range of symbols is shown in the annex.

8.2 The equivalence relationship

8.2.1 General

This is the relationship between preferred and non-preferred terms where two or more terms are regarded, for indexing purposes, as referring to the same concept. In a printed thesaurus, preferred and non-preferred terms should be distinguished typographically whenever possible. Reciprocity is expressed by the following conventions:

USE, written as a prefix to the preferred term;

UF (use for), written as a prefix to the non-preferred term.

Example:

BIRDS
  UF Aves

Aves
  USE BIRDS

This general relationship covers two types of term:

a) synonyms;

b) quasi-synonyms.

8.2.2 Synonyms

Synonyms are terms whose meanings can be regarded as the same in a wide range of contexts, so that they are virtually interchangeable. They occur more frequently in a controlled indexing language, where meanings are deliberately restricted, than in natural language. Various types of synonym are encountered in practice. The following list is not intended to be exhaustive, but it indicates some of the more common classes of synonyms encountered in practice:

a) terms of different linguistic origin

Example

POLYGLOT; MULTILINGUAL

b) popular names and scientific names

Examples:

PENGUINS; SPHENISCIFORMES

ASPIRIN; ACETYLSALICYLIC ACID

c) common nouns and trade names

Example:

VACUUM FLASKS; THERMOS FLASKS

d) variant names for emergent concepts

Example:

HOVERCRAFT; AIR CUSHION VEHICLES

e) current or favoured terms versus outdated or deprecated terms

Examples:

RADIO; WIRELESS

DEVELOPING COUNTRIES; UNDERDEVELOPED COUNTRIES

f) variant spellings, including stem variants and irregular plurals

Examples:

ROMANIA; RUMANIA; ROUMANIA

GEESE; GOOSE

g) terms originating from different cultures sharing a common language

Examples:

FLATS; APARTMENTS

LIFTS; ELEVATORS

h) abbreviations and full names

Example:

PVC; POLYVINYL CHLORIDE

i) the factored and unfactored form of a compound term

Example:

COAL + MINING; COAL MINING

In these and similar cases, preferred terms should be selected to serve the needs of the majority of users, bearing in mind the criteria proposed in clause 6. For the sake of predictability, these criteria should be applied consistently throughout the thesaurus. If, for example, it is decided that popular names rather than scientific names should serve as preferred terms, the indexer should adhere consistently to this decision except when an appropriate popular name does not exist.

8.2.3 Quasi-synonyms

Quasi-synonyms are terms whose meanings are generally regarded as different in ordinary usage, but they are treated as though they are synonyms for indexing purposes. These terms frequently represent points on a continuum:

Example:

WETNESS, DRYNESS

The extent to which terms are treated as quasi-synonyms depends to a large extent upon the subject field covered by the thesaurus. An indexing agency concerned with "washing equipment", for example, may decide to designate either "wetness" or "dryness" as the preferred term when dealing with documents on either of these concepts, on the understanding that a specialist studying one of these properties should also retrieve all documents on the other. This situation will not necessarily apply in a different subject field, and quasi-synonyms are likely to be least common in an indexing language which covers several disciplines. As a general rule terms should be treated as quasi-synonyms only in fringe subject areas. They should not be used as a means for reducing the number of preferred terms in an indexing language.

8.2.4 Upward posting

This refers to a technique in which the name of a class and also the names of its members are treated as an equivalence set, the broader term then fuctioning as the preferred term.

Example:

ROCKS
  UF basalt
    granite
    slate
    etc.

basalt **USE** ROCKS

granite **USE** ROCKS

slate **USE** ROCKS

This technique, which is sometimes used to reduce the number of terms in an indexing language, should generally be avoided. If employed, it should be applied only to terms in the fringe area of the subject field covered by the thesaurus.

8.3 The hierarchical relationship

8.3.1 General

This basic relationship most distinguishes a systematic thesaurus from an unstructured list of terms, for example a glossary or dictionary. It is based on degrees or levels of superordination and subordination, where the superordinate term represents a class or whole, and subordinate terms refer to its members or parts. Reciprocity is expressed by the following abbreviations:

BT (i.e. broader term), written as a prefix to the superordinate term;

NT (i.e. narrower term), written as a prefix to the subordinate term.

Example:

AIRLINES
BT TRANSPORT SERVICES

TRANSPORT SERVICES
NT AIRLINES

8.3.2 Kinds of hierarchical relationship

This general relationship covers three logically different situations distinguished as follows:

a) the generic relationship;

b) the hierarchical whole-part relationship;

c) the instance relationship.

Each of these leads to hierarchies which are amenable to a logical test through reference to the basic types of concept represented by the terms, for example those listed in 6.1.1. Every subordinate term should refer to the same basic kind of concept as its superordinate term; that is, both the broader and narrower term should represent a thing, or an action, or a property, etc.

Example:

a) METALS (a class of materials) and CASTING (an action) represent different types of concept, and these terms cannot therefore be related hierarchically;

b) METALS and NON-FERROUS METALS both represent materials; these terms could, therefore, be generically related.

8.3.3 Use of node labels

When terms are arranged as hierarchies in a systematic thesaurus (see 9.3), a node label is sometimes inserted into the hierarchy to show which characteristic was used to divide the class (represented by the broader term) into its species (represented by narrower terms).

Example:

CARS

By motive power

DIESEL CARS

ELECTRIC CARS

By purpose

RACING CARS

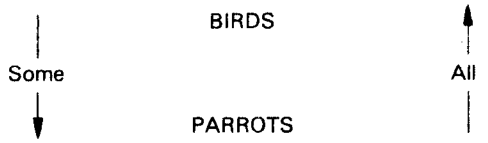
SPORTS CARS

— where **By motive power** and **By purpose** function as node labels.

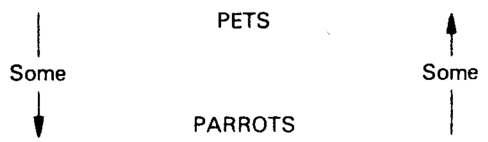
The use of node labels is described more fully in 9.3.3.

8.3.4 The generic relationship

8.3.4.1 This relationship identifies the link between a class or category and its members or species. In addition to the test for validity described in 8.3.2, this relationship is also amenable to a logical "all-and-some" test, as shown in the following diagram:



The diagram above indicates that some members of the class "birds" are known as "parrots", and all "parrots", by definition and regardless of context, are regarded as "birds". This test would usually ensure that a term such as "parrots" is not subordinated to a class such as "pets", since parrots are not pets by definition. This particular relational link could be represented by the diagram:



— that is, some members of the class "pets" are known as "parrots", but only some parrots are regarded as pets. These terms should therefore be assigned to different categories in the thesaurus, and both should be assigned to the same document when indexing a work on "parrots as pets". This test is intended to avoid subjectivity when assigning terms to categories. It would not, however, offer a reasonable basis for assigning terms such as "pets" and "parrots" to different categories in a specialist thesaurus devoted to domestic animals, where parrots are regarded as a kind of pet according to the frames of reference of the majority of users. In that case, "parrots" could be subordinated to "pets" in the same hierarchy. It needs to be recognized, however, that this is a special case.

8.3.4.2 If it is considered necessary, the generic relationship can be identified by the following abbreviations or their equivalents in other languages:

BTG: Broader term (generic)

NTG: Narrower term (generic)

Example:

RATS
  BTG RODENTS

RODENTS
  NTG RATS

8.3.5 The hierarchical whole-part relationship

8.3.5.1 This relationship covers a limited range of situations where the name of a part implies the name of its possessing whole in any context. The terms can then be organized as a hierarchy, the name of the whole serving as the superordinate term, and the name of the part as the subordinate term. This applies to four main classes of terms:

a) systems and organs of the body

Example:

CIRCULATORY SYSTEM
  CARDIO-VASCULAR SYSTEM
    ARTERIES
    VEINS

b) geographical locations

Example:

CANADA
  ONTARIO
    OTTAWA
    TORONTO

c) disciplines or fields of discourse

Example:

SCIENCE
  BIOLOGY
    BOTANY
    ZOOLOGY

d) hierarchical social structures

Example:

ARMIES
  CORPS
    DIVISIONS
      BATTALIONS
      REGIMENTS

8.3.5.2 In cases such as those listed as a) to d) above, the hierarchy is based upon a generally recognized whole-part relationship which could be established by checking standard reference works, for example dictionaries or gazetteers. Other kinds of term may also be organized into hierarchies on the basis of their whole-part relationships in thesauri which are limited to special subject fields. For example, a thesaurus devoted to "turbine engineering" may contain a hierarchy such as

TURBINES
  COMPRESSORS
    BLADES

This should not be attempted, however, in a thesaurus outside the limited field in which the name of the whole is implied by the name of the part. In a general thesaurus, for example, a term such as "blades" can occur in a number of different contexts. In these cases, the name of the whole and its part(s) should be assigned to different categories in the thesaurus, and both terms should be assigned to a document on these concepts.

8.3.5.3 If it is considered necessary, the hierarchical whole-part relationship can be indicated specifically by the following abbreviations or their equivalents in other languages:

BTP = broader term (partitive)

NTP = Narrower term (partitive)

Examples:

CENTRAL NERVOUS SYSTEM
  BTP NERVOUS SYSTEM

NERVOUS SYSTEM
  NTP CENTRAL NERVOUS SYSTEM

8.3.6 The instance relationship

8.3.6.1 This relationship identifies the link between a general category of things or events, expressed by a common noun, and an individual instance of that category, the instance then forming a class-of-one which is represented by a proper name.

Example:

MOUNTAIN REGIONS — Class

ALPS
HIMALAYAS } — Instances

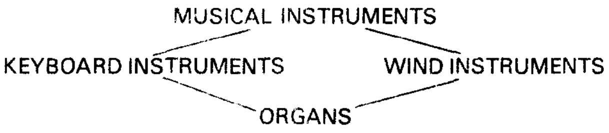
In this example, the "Alps" and "Himalayas" are assigned to subordinate positions in a hierarchy, yet they are neither kinds nor parts of "mountain regions", but represent specific examples or instances.

8.3.6.2 As noted earlier, proper names are frequently excluded from thesauri, on the grounds that, if admitted, they would overload the categories. If these terms are excluded, the name of the instance and also its broader term should both be assigned to any document that deals with the instance.

8.3.7 Polyhierarchical relationships

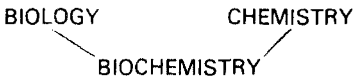
Some concepts can belong, on logical grounds, to more than one category at the same time. They are then said to possess polyhierarchical relationships.

Example:



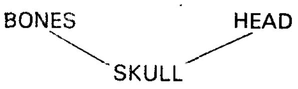
In this example, the term "organs" is assigned to subordinate positions on the basis of its generic relationship to two broader terms. In other cases, polyhierarchical links may be based upon whole-part relationships.

Example:



In some cases, polyhierarchical links can be based upon logically different relationships.

Example:



In this example, the link between "bones" and "skull" is based upon the generic relationship (the skull is a kind of bone), whereas the link between "head" and "skull" is based upon the hierarchical whole-part relationship (the skull is a part of the head).

8.4 The associative relationship

8.4.1 General

This basic relationship is easier to define in terms of negative rather than positive characteristics. It covers relationships between pairs of terms which are not members of an equivalence set, or can they be organized as a hierarchy in which one term is subordinated to another, yet they are mentally associated to such an extent that the link between them should be made explicit in the thesaurus, on the grounds that it would reveal alternative terms which might be used for indexing or retrieval. This relationship is reciprocal, and is indicated by the abbreviation "**RT**" (related term), or its equivalent in other languages.

Example:

BIRDS
  RT ORNITHOLOGY

ORNITHOLOGY
  RT BIRDS

It is important to exercise strict control over the choice of terms linked in this way, and to avoid subjective judgements. As a general guideline it can be stated that one of the terms should be strongly implied, according to the frames of reference shared by the users of an index, whenever the other is employed as an indexing term. More specifically, it will frequently be found that one of the terms is a necessary component in any explanation or definition of the other, to the extent that the term "birds", for example, forms a necessary part of the explanation of "ornithology".

Two kinds of term can be linked by the associative relationship:

a) those that belong to the same category;

b) those belonging to different categories.

These are considered separately below.

8.4.2 Terms belonging to the same category

8.4.2.1 This relates to siblings with overlapping meanings, such as "ships" and "boats", where each of the terms is amenable to an exact definition, (consequently they do not form an equivalence set), yet they are sometimes used loosely and almost interchangeably, so that the user seeking documents on one of the terms should be reminded of the other. This link does not need to be made explicit in the systematic section of a thesaurus containing organized hierarchies (see 9.3 below), since this form of display automatically brings such terms together. The link should be indicated, however, in an alphabetical thesaurus, and in the alphabetical section of a systematic thesaurus.

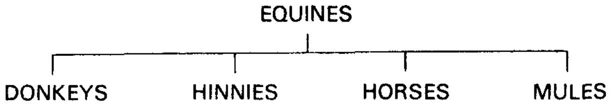
Example:

BOATS
  BT VEHICLES
  RT SHIPS

SHIPS
BT VEHICLES
RT BOATS

8.4.2.2 It is not necessary to interrelate all sibling terms in this way. For example, there is no need to associate such terms as "horses" and "donkeys" on the grounds that they share a common broader term, i.e. "equines", since the meaning of the terms do not overlap in this case.

8.4.2.3 Concepts linked by a familial or derivational relationship (i.e. one of the concepts was derived from the other), can also be regarded as belonging to this group. This would apply to terms such as "hinnies" and "mules", which represent kinds of crossbreed between "horses" and "donkeys". The relationships between these concepts can be represented by a tree diagram.



These interrelationships could be expressed as follows in the alphabetical section of a thesaurus:

EQUINES
NT DONKEYS
NT HINNIES
NT HORSES
NT MULES

DONKEYS
BT EQUINES
RT HINNIES
RT MULES

HINNIES
BT EQUINES
RT DONKEYS
RT HORSES
RT MULES

HORSES
BT EQUINES
RT HINNIES
RT MULES

MULES
BT EQUINES
RT DONKEYS
RT HINNIES
RT HORSES

8.4.3 Terms belonging to different categories

Many grounds can be established for associating terms belonging to different categories (i.e. they refer to different conceptual types), while satisfying the requirements that one of the terms should be strongly implied by the other. The following groups are offered only as representative examples of typical relational situations encountered in practice:

a) a discipline or field of study and the objects or phenomena studied

Examples:

1) FORESTRY
RT FORESTS

FORESTS
RT FORESTRY

2) AESTHETICS
RT BEAUTY

BEAUTY
RT AESTHETICS

b) an operation or process and its agent or instrument

Examples:

1) TEMPERATURE CONTROL
RT THERMOSTATS

THERMOSTATS
RT TEMPERATURE CONTROL

2) DATA PROCESSING
RT COMPUTER SYSTEMS

COMPUTER SYSTEMS
RT DATA PROCESSING

c) an action and the product of the action

Examples:

1) WEAVING
RT CLOTH

CLOTH
RT WEAVING

2) CONSTRUCTION
RT STRUCTURES

STRUCTURES
RT CONSTRUCTION

d) an action and its patient

Examples:

1) HARVESTING
RT CROPS

CROPS
RT HARVESTING

2) IMPRISONMENT
RT PRISONERS

PRISONERS
RT IMPRISONMENT

e) concepts related to their properties

Examples:

1) POISONS
RT TOXICITY

TOXICITY
RT POISONS

2) PERCEPTION
RT ACUITY

ACUITY
RT PERCEPTION

f) concepts related to their origins

Example:

DUTCH
RT NETHERLANDS

NETHERLANDS
RT DUTCH

g) concepts linked by causal dependence

Examples:

1) BEREAVEMENT
RT DEATH

DEATH
RT BEREAVEMENT

2) DISEASES
RT PATHOGENS

PATHOGENS
RT DISEASES

h) a thing and its counter agent

Examples:

1) PLANTS
RT HERBICIDES

HERBICIDES
RT PLANTS

2) AIRCRAFT
RT ANTI-AIRCRAFT WEAPONS

ANTI-AIRCRAFT WEAPONS
RT AIRCRAFT

i) a concept and its unit of measurement

Example:

ELECTRIC CURRENT
RT AMPERE

AMPERE
RT ELECTRIC CURRENT

j) syncategorematic phrases and their embedded nouns

Examples:

1) MODEL SHIPS
RT SHIPS

SHIPS
RT MODEL SHIPS

2) FOSSIL REPTILES
RT REPTILES

REPTILES
RT FOSSIL REPTILES

8.4.4 Node labels

In order to bring closely-related concepts together within a systematic thesaurus (see 9.3), related terms are sometimes printed in the same schedule as the terms with which they are generally associated. The related terms (which do not form a logical part of the hierarchy) should then be identified by a node label.

Example:

BOOKS

Operations

BINDING
PRINTING

— where **Operations** functions as a node label.

The use of node labels is described more fully in 9.3.3.

9 Display of terms and their relationships

9.1 General

Terms and their interrelationships can be displayed in various ways in a thesaurus. This International Standard does not set out to review all possible forms of display, but is limited to descriptions of three basic methods in common use

a) alphabetical display, with scope notes and indications of inter-term relationships;

b) systematic display, supported by an alphabetical index;

c) graphic display, with an alphabetical section.

Existing thesauri usually belong to one of these three basic types, although a given thesaurus may include features of more than one. For example, a thesaurus in which the main part is organized systematically may be supported by an alphabetical index which also indicates a relatively wide range of inter-term relationships. In thesauri consisting of more than one part (for example those with systematic or graphic displays), all information relating to each term should be brought together in one of the parts.

These basic types are described in turn below, and each is illustrated by a typical arrangement of a common set of terms related to "Cameras".

9.2 Alphabetical display

9.2.1 In this form of thesaurus all indexing terms, whether preferred or non-preferred, are organized as a single alphabetical sequence. Non-preferred terms are usually accompanied only by references (for example **USE**) to their preferred equivalents. Where ancillary information is appended to preferred terms it should be listed in the following order:

1) **SN** scope notes or definitions.

2) **UF** references to non-preferred equivalent term(s).

3) **TT** references to top terms, if necessary.

4) **BT** references to broader terms.

5) **NT** references to narrower terms.

6) **RT** references to related terms.

9.2.2 An example of this form of display is shown in figure 1. This is probably the easiest type of thesaurus to construct and reproduce, but it should be borne in mind that the user of an alphabetical thesaurus cannot survey, at a single position, all the broader and narrower terms that together constitute a hierarchy. Extra relational information can be added to an alphabetical display to indicate, for example, the top term in the hierarchy to which a specific concept belongs (using the abbreviation **TT**).

Example:

UNDERWATER CAMERAS
  TT OPTICAL EQUIPMENT
  BT CAMERAS
  NT UNDERWATER CINE CAMERAS
  RT DIVING

9.2.3 It may be useful to indicate more than one level of subordination and/or superordination, using abbreviations such as

BT1 = Broader term on first (i.e. next) level of supeordination.

BT2 = Broader term on second (i.e. higher) level of superordination.

NT1 = Narrower term on first (i.e. next) level of subordination.

NT2 = Narrower term on second (i.e. lower) level of subordination.

This could also be indicated by different levels of indention.

9.3 Systematic display

9.3.1 Parts of a systematic thesaurus

9.3.1.1 A thesaurus in which terms are organized systematically should consist of two parts:

a) categories or hierarchies of terms arranged according to their meanings and logical interrelationships;

b) an alphabetical index which directs the user to the appropriate part(s) of the systematic section.

The link between these two sections is provided by a system of addresses. An address code is assigned to each of the preferred terms in the systematic section, and this code functions as a reference in the alphabetical index. These codes should have obvious filing values. They may consist simply of running numbers [as in the example in figures 2a) and 2b)], or they may comprise a system of hierarchically expressive notation.

9.3.1.2 With this type of display the systematic section is frequently regarded as the main part of the thesaurus, i.e. the part which carries most of the definitional and relational information, in which case the alphabetical index assumes the role of a complementary but secondary component. This is not invariable, however, and thesauri differ widely in their approaches to the relative importance and functions of these two sections, and also in the arrangement and kind of relational information provided in each. Terms in the alphabetical section are sometimes accompanied by a level of relational and other information such that the two parts can be regarded as having equal status.

35 mm CAMERAS
  BT MINIATURE CAMERAS

CAMERAS
  BT OPTICAL EQUIPMENT
  NT MOVING PICTURE CAMERAS
    STEREO CAMERAS
    STILL CAMERAS
    UNDERWATER CAMERAS
  RT PHOTOGRAPHY

CINE CAMERAS
  BT MOVING PICTURE CAMERAS
  NT UNDERWATER CINE CAMERAS
  RT CINEMA

CINEMA
  RT CINE CAMERAS

DIVING
  RT UNDERWATER CAMERAS

INSTANT PICTURE CAMERAS
  SN Cameras which produce a finished print directly
  BT STILL CAMERAS

Land cameras **USE** VIEW CAMERAS

MICROSCOPES
  BT OPTICAL EQUIPMENT

MINIATURE CAMERAS
  BT STILL CAMERAS
  NT 35 mm CAMERAS

MOVING PICTURE CAMERAS
  BT CAMERAS
  NT CINE CAMERAS
    TELEVISION CAMERAS

OPTICAL EQUIPMENT
  NT CAMERAS
    MICROSCOPES

PHOTOGRAPHY
  RT CAMERAS

REFLEX CAMERAS
  BT STILL CAMERAS
  NT SINGLE LENS REFLEX CAMERAS
    TWIN LENS REFLEX CAMERAS

SINGLE LENS REFLEX CAMERAS
  BT REFLEX CAMERAS

STEREO CAMERAS
  BT CAMERAS

STILL CAMERAS
  BT CAMERAS
  NT INSTANT PICTURE CAMERAS
    MINIATURE CAMERAS
    REFLEX CAMERAS
    VIEW CAMERAS

TELEVISION
  RT TELEVISION CAMERAS

TELEVISION CAMERAS
  SN Cameras for both broadcasting and videotape recording
  BT MOVING PICTURE CAMERAS
  RT TELEVISION

TWIN LENS REFLEX CAMERAS
  BT REFLEX CAMERAS

UNDERWATER CAMERAS
  BT CAMERAS
  NT UNDERWATER CINE CAMERAS
  RT DIVING

UNDERWATER CINE CAMERAS
  BT CINE CAMERAS
    UNDERWATER CAMERAS

VIEW CAMERAS
  SN Cameras with through-the-lens focusing and a range of movements of the lens plane relative to the film plane
  UF land cameras
  BT STILL CAMERAS

NOTE — No attempt has been made to collect and organize all relevant terms within this field; terms have been selected primarily to demonstrate thesaural features considered in earlier clauses, for example the treatment of scope notes, synonyms, broader and narrower terms, etc.

Figure 1 — Alphabetical display

301 OPTICAL EQUIPMENT

302 CAMERAS
RT PHOTOGRAPHY 824

303 MOVING PICTURE CAMERAS
By medium

304 CINE CAMERAS
RT CINEMA 895

305 UNDERWATER CINE CAMERAS

306 TELEVISION CAMERAS
RT TELEVISION 897

307 STEREO CAMERAS

308 STILL CAMERAS

309 INSTANT PICTURE CAMERAS
SN Cameras which produce a finished print directly

310 MINIATURE CAMERAS

311 35 mm CAMERAS

312 REFLEX CAMERAS

313 SINGLE LENS REFLEX CAMERAS

314 TWIN LENS REFLEX CAMERAS

315 VIEW CAMERAS
SN Cameras with through-the-lens focusing and a range of movements of the lens plane relative to the film plane
UF land cameras

316 UNDERWATER CAMERAS
RT DIVING 931

317 UNDERWATER CINE CAMERAS

318 MICROSCOPES

Figure 2a) — Systematic display

35 mm CAMERAS	311
CAMERAS	302
RT PHOTOGRAPHY	824
CINE CAMERAS	304
RT CINEMA	895
CINEMA	895
RT CINE CAMERAS	304
DIVING	931
RT UNDERWATER CAMERAS	316
INSTANT PICTURE CAMERAS	309
SN Cameras which produce a finished print directly	
land cameras **USE** VIEW CAMERAS	315
MICROSCOPES	318
MINIATURE CAMERAS	310
MOVING PICTURE CAMERAS	303
OPTICAL EQUIPMENT	301
PHOTOGRAPHY	824
RT CAMERAS	302
REFLEX CAMERAS	312
SINGLE LENS REFLEX CAMERAS	313
STEREO CAMERAS	307
STILL CAMERAS	308
TELEVISION	897
RT TELEVISION CAMERAS	306
TELEVISION CAMERAS	306
RT TELEVISION	897
TWIN LENS REFLEX CAMERAS	314
UNDERWATER CAMERAS	316
RT DIVING	931
UNDERWATER CINE CAMERAS	305; 317
VIEW CAMERAS	315
SN Cameras with through-the-lens focusing and a range of movements of the lens plane relative to the film plane	
UF land cameras	

Figure 2b) — Alphabetical index to systematic display

9.3.1.3 Typical layouts are illustrated in figures 2a) (systematic section) and 2b) (alphabetical index). Attention is drawn to the following features:

a) The systematic section [figure 2a)] contains scope notes and references to equivalent and related terms. Explicit references to broader and narrower terms are not attached to each of the terms, since these relationships are sufficiently indicated by the position of each term within its hierarchy, and by its level of indention. The following features should also be noted:

1) related terms are accompanied by address codes indicating their positions in other hierarchies;

2) node labels (for example **By medium**) are occasionally inserted to show the basis on which a category has been divided (see also 9.3.3);

b) The alphabetical section [figure 2b)] also contains scope notes and reciprocal references to synonyms and related terms. It does not, in this instance, contain explicit references to broader and narrower terms, but these hierarchical relationships could have been indicated in the alphabetical section.

9.3.2 Primary organization of a systematic thesaurus

9.3.2.1 General

When terms are organized into hierarchies or categories in the systematic section, the compiler needs to consider not only the inter-term relationships described in 8.2 to 8.4, but also relationships between the hierarchies or categories themselves. In effect, it is necessary to impose an overall structure or macroclassification on the systematic section to ensure that similar concepts are brought together and unrelated concepts are separated for the benefit of users. Two general approaches can be adopted to the primary organization of a systematic thesaurus:

a) organization into fields or disciplines;

b) organization by facets.

These approaches are considered separately in 9.3.2.2 and 9.3.2.3 where it will be see that each possesses its advantages and disadvantages. In practice these techniques are frequently combined: that is to say, a thesaurus may be organized primarily into fields, and facets may then be introduced as the basis for organizing concepts within a given field. This combined approach is considered in 9.3.2.4.

9.3.2.2 Organization into fields

This approach is basically similar to that adopted by the compiler of a library classification who begins by organizing the universe of subjects into main classes or disciplines. When applied to a thesaurus, categories of concepts are grouped initially to reflect the various fields of interests of its users. This technique is clearly applicable to a thesaurus covering a range of different disciplines. It would then ensure that terms which are usually associated with a given field, such as "art", are brought together, and are also effectively separated from those belonging to different areas of interest, such as "economics" or "physics". This approach can also be applied to a thesaurus which is ostensibly restricted to a single field. A thesaurus devoted to "medicine", for example, which might be organized initially into subfields such as "injuries", "diseases", "treatments", etc., is also likely to include a number of terms from different disciplines, for example "management", "law" and "data processing", which should be separated from the medical terms. Although this approach ensures that concepts are organized into groups which generally correspond to the modes of thinking of its users, it also possesses an inherent disadvantage. A given concept may be assigned to different categories in different thesauri, depending upon their grouping by field of interest. This could introduce problems when information is exchanged between different documentation centres. Extensive revision may also be necessary when changes occur in the boundaries of subject fields, or when new disciplines emerge.

9.3.2.3 Organization by facets

The notion of a facet is more abstract than the concept of a field, and organization by facets involves a more rigorous mental discipline on the part of the thesaurus compiler. When applying this technique, terms are organized into classes or sets according to the basic kinds of concept they represent, with no initial regard for the field or fields with which a given concept is usually associated. For this purpose, concepts should generally be distinguished at a higher level of specificity than that seen in the outline classes of conceptual types listed in 6.1.1. Concrete things, for example, might be divided into two main categories, i.e. "naturally occurring things" and "artificial things"; the latter might be further distinguished as "products", "tools", etc., depending upon the subject field. Actions are frequently subdivided into "operations" (i.e. transitive actions carried out upon things), and "processes" (these can generally be regarded as intransitive actions initiated by or occurring within things).

Organization of terms into facets offers certain advantages:

a) the need for a major revision is less likely to occur if a given concept changes the field with which it is usually associated;

b) a higher level of agreement between different documentation centres could be expected.

It is also necessary, however, to note the following disadvantages:

a) organization by facets tends to scatter the concepts usually associated with a given field or discipline;

b) the basis on which the thesaurus is organized will be generally less self-evident to both indexers and users.

9.3.2.4 Combined approach.

In practice, the different approaches described in 9.3.2.2 and 9.3.2.3 are frequently combined, for example when a thesaurus organized primarily into subject fields is further subdivided by facets.

Example: A field such as "education" may be divided into facets representing basic categories such as **operations** (for example, "teaching"), **processes** (for example, "learning"), and **agents** (for example, "teachers").

9.3.3 Node labels

Two kinds of node label, each with a different function, may be inserted into the schedules of a systematic thesaurus:

a) those that serve as facet indicators to show the logical basis on which the hierarchy was organized, where the terms on either side of the node label refer to concepts of the same basic type.

Example:

AIRCRAFT
By payload
FREIGHT AIRCRAFT
PASSENGER AIRCRAFT
By user
CIVIL AIRCRAFT
MILITARY AIRCRAFT

In this example, **By plyaload** and **By user** function as node labels. They introduce terms that are narrower than the class preceding the node label.

b) labels used to introduce different types of concepts, ensuring that these are collocated, for the convenience of users, under the concept(s) with which they are usually associated.

Example:

BOOKS
Operations
BINDING
PRINTING
etc.

Operations functions as the node label. In these cases, the apparently "broader" term, for example "books", serves only to collocate the associated concepts introduced by the node label. In the alphabetical section, these terms should be indicated as associated terms (**RT**), not as broader and narrower terms (**BT/NT**).

Example:

BOOKS		BOOKS
RT BINDING	but **NOT**	**NT** BINDING
BINDING **RT** BOOKS		BINDING **BT** BOOKS

9.3.4 Polyhierarchical relationships (see also 8.3.7)

9.3.4.1 The standard treatment of polyhierarchical terms is demonstrated in figures 2a) and 2b), where the term "UNDERWATER CINE CAMERAS" belongs to two classes, "304 CINE CAMERAS" and "316 UNDERWATER CAMERAS", and therefore appears at two positions in the systematic section. The polyhierarchical term is then accompanied by both addresses in the alphabetical section.

9.3.4.2 A different approach may be adopted in circumstances such as the following:

a) space in the systematic section is strictly limited;

b) an extensive entry is needed for the polyhierarchical term. It might be accompanied, for example, by a long scope note, and/or several non-preferred synonyms, narrower terms and related terms.

It may then be decided to print the full entry under only one of the broader terms to which the polyhierarchical term logically belongs, and make references to this entry not only from the alphabetical section, but also (as an added **NT**) from each of the other broader terms in the systematic section. A note of each of these broader terms should be incorporated (as an added **BT**) into the full entry.

Example (assuming that "UNDERWATER CINE CAMERAS" calls for an extensive entry at position 317):

304 CINE CAMERAS
Extra **NT** UNDERWATER CINE CAMERAS 317
RT CINEMA 895
....
....

316 UNDERWATER CAMERAS

317 UNDERWATER CINE CAMERAS
Extra **BT** CINE CAMERAS 304
....
....

9.4 Graphic display

9.4.1 In this type of display, indexing terms and their interrelationships are set out in the form of a two-dimensional figure which allows the indexer or user to survey a range of interrelated terms. Various kinds of graphic display are found in published thesauri, but two main types can be identified:

a) tree structures;

b) arrowgraphs.

9.4.2 A thesaurus which incorporates a graphic display should consist of two parts:

a) The graphic display

This is usually limited to preferred terms only, since this format does not readily accommodate scope notes, synonyms, etc. The main display is identified by a symbol, for example a running number or an element of hierarchically expressive notation. Symbols need not be assigned to individual terms, although the position of a term within the figure can be indicated by the use of a grid system as described below. The symbols used to identify displays also function as addresses in the alphabetical index.

b) The alphabetical index

This contains a full range of scope notes and references between preferred and non-preferred terms, and it can also include references to broader, narrower and relative terms. The indexes used as examples [i.e. figures 3b) and 4b)] differ in the extent to which they refer to broader and narrower terms:

1) the index to the tree structure [figure 3b)] does not indicate hierarchical relationships;

2) the index to the arrowgraph [figure 4b)] includes not only scope notes etc., but it also contains references to broader and narrower terms.

In cases where the alphabetical section contains the higher proportion of definitional and relational information, it functions as the main part of the thesaurus. It may be as full as the alphabetical thesaurus described in 9.2, in which case the graphic display should be seen as having the supportive role. In thesauri covering special subject fields, graphic displays are frequently limited to terms in the core subjects, whereas all terms appear in the alphabetical section.

9.4.3 A tree structure is illustrated in figure 3a). "Cameras" appears as the broadest term at the head of the diagram, and narrower terms are printed in subordinate positions, with hierarchical relationships indicated by the connecting lines. This kind of display is more effective when the layout is relatively simple. Consequently, scope notes, synonyms and related terms have been excluded from the tree diagram, and are assigned instead to the alphabetical section. The tree as a whole is identified by the symbol E417, and this is appended, as an address, to each of its terms in the alphabetical section shown in figure 3b). The alphabetical section also contains terms which belong to other (unillustrated) trees, for example "television", which belongs to a structure labelled as R685.

9.4.4 An arrowgraph is illustrated in figure 4a), where the broadest term "cameras" is placed in a relatively central position and is typographically emphasized. Narrower terms are printed in other positions within the grid, and levels of subordination are indicated by the connecting lines and the relative distance of the subordinate term from the broadest term. Related terms, each accompanied by an address indicating the network in which it occurs, are printed outside the main grid, the connections being shown by pecked lines. The whole display is identified by the symbol E 417, and this forms part of the address of each term in the alphabetical section reproduced as figure 4b). In addition, the position of each individual term within the figure is identified through reference to a grid system, and these grid references are also printed in the alphabetical section. In this example, the alphabetical section contains a full range of definitional and relational information, and it could serve as an alphabetical thesaurus in its own right.

E417

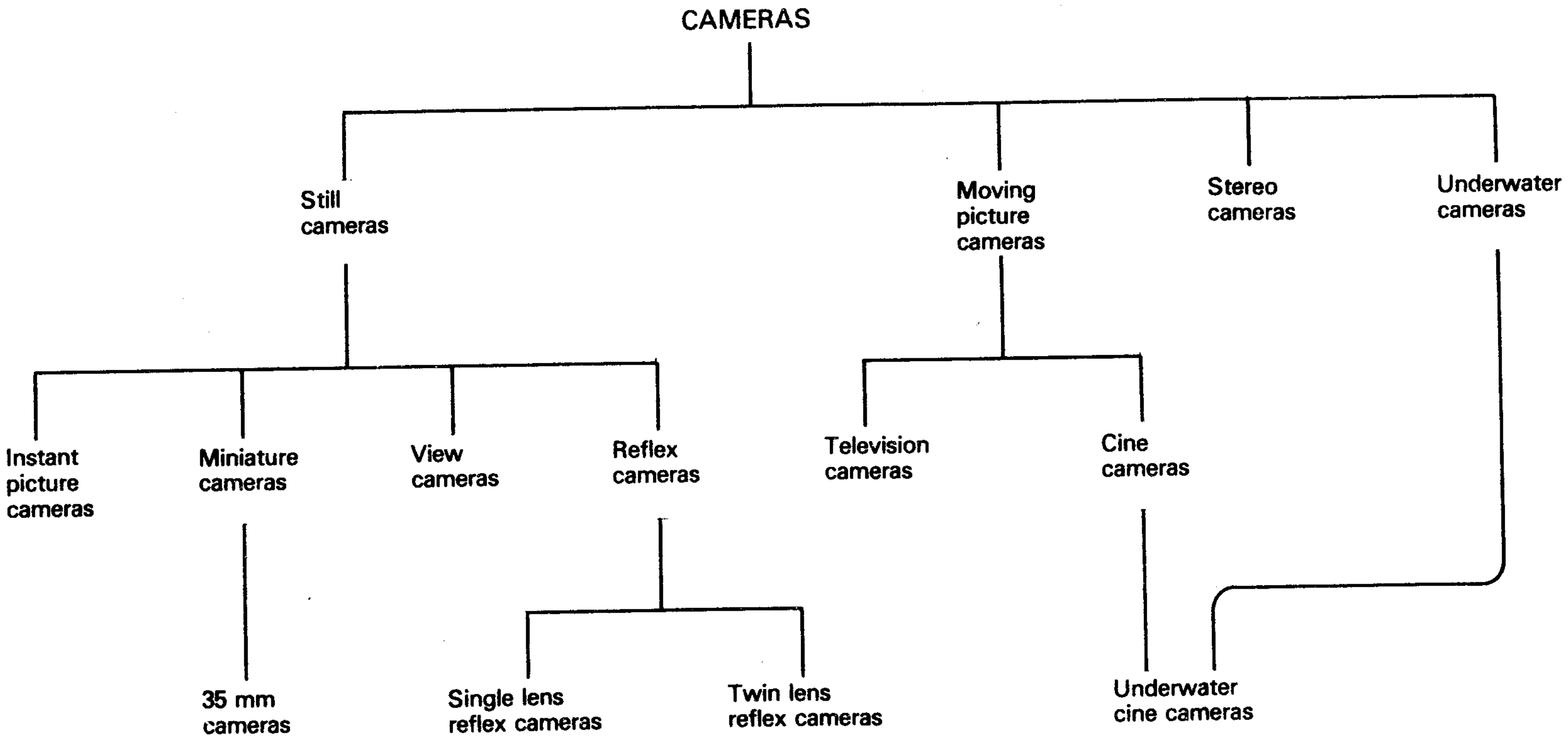


Figure 3a) — Tree structure

35 mm CAMERAS	E 417
CAMERAS	E 417
RT PHOTOGRAPHY	R 562
CINE CAMERAS	E 417
RT CINEMA	R 668
CINEMA	R 668
RT CINE CAMERAS	E 417
DIVING	T 473
RT UNDERWATER CAMERAS	E 417
INSTANT PICTURE CAMERAS	E 417
SN Cameras which produce a finished print directly	
land cameras **USE** VIEW CAMERAS	
MINIATURE CAMERAS	E 417
MOVING PICTURE CAMERAS	E 417
PHOTOGRAPHY	R 562
RT CAMERAS	E 417
REFLEX CAMERAS	E 417
SINGLE LENS REFLEX CAMERAS	E 417
STEREO CAMERAS	E 417
STILL CAMERAS	E 417
TELEVISION	R 685
RT TELEVISION CAMERAS	E 417
TELEVISION CAMERAS	E 417
RT TELEVISION	R 685
TWIN LENS REFLEX CAMERAS	E 417
UNDERWATER CAMERAS	E 417
RT DIVING	T 473
UNDERWATER CINE CAMERAS	E 417
VIEW CAMERAS	E 417
SN Cameras with through-the-lens focusing and a range of movements of the lens plane relative to the film plane	
UF land cameras	

Figure 3b) — Alphabetical index to tree structure

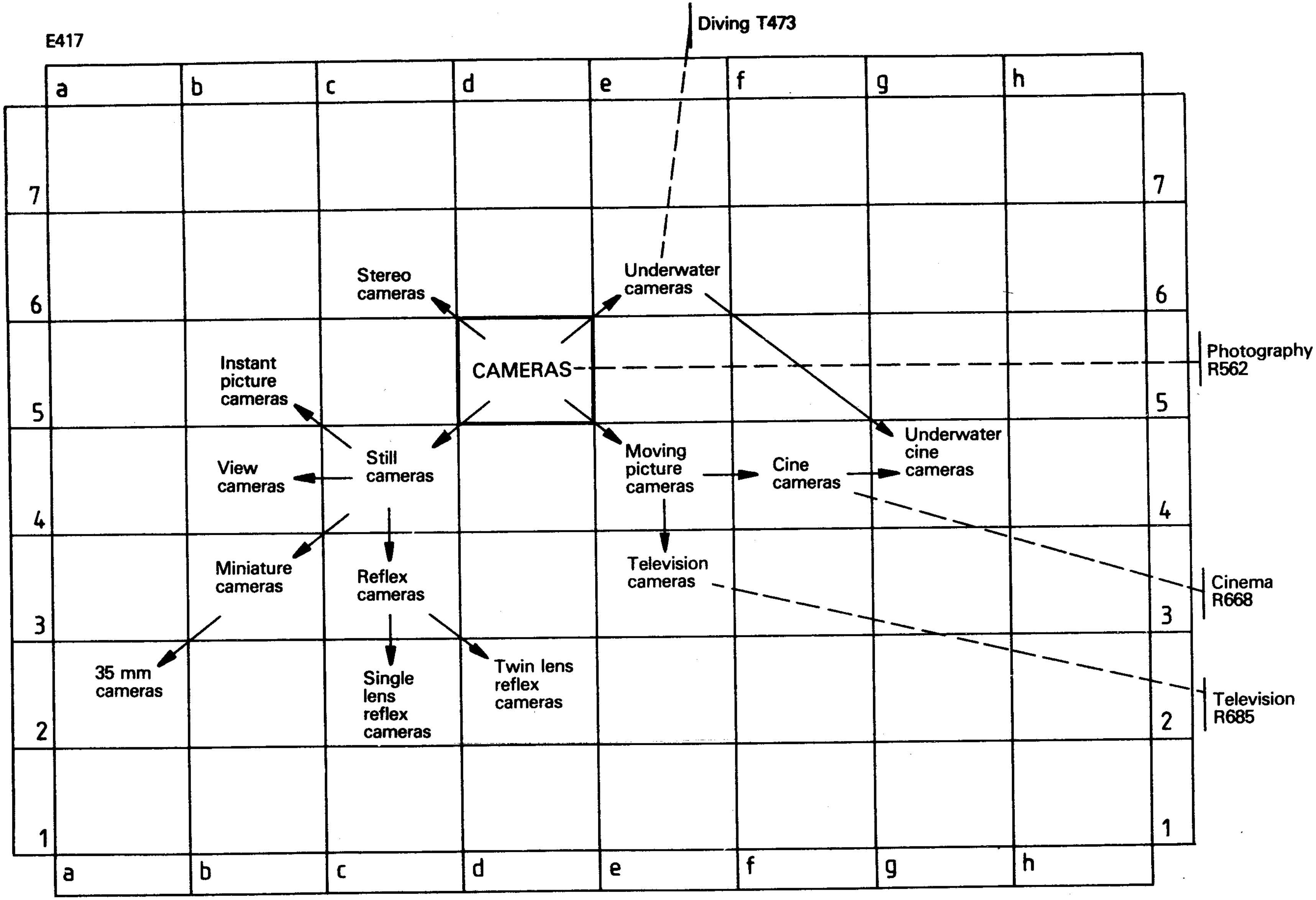


Figure 4a) — Arrowgraph

35 mm CAMERAS E 417.a2
BT MINIATURE CAMERAS

CAMERAS E 417.d5
RT PHOTOGRAPHY R 562
NT MOVING PICTURE CAMERAS
STEREO CAMERAS
STILL CAMERAS
UNDERWATER CAMERAS

CINE CAMERAS E 417.f4
BT MOVING PICTURE CAMERAS
NT UNDERWATER CINE CAMERAS
RT CINEMA R 668

CINEMA R 668.d5
RT CINE CAMERAS E 417

DIVING T 473.g5
RT UNDERWATER CAMERAS E 417

INSTANT PICTURE CAMERAS E 417.b5
SN Cameras which produce a finished print directly
BT STILL CAMERAS

land cameras **USE** VIEW CAMERAS

MINIATURE CAMERAS E 417.b3
BT STILL CAMERAS
NT 35 mm CAMERAS

MOVING PICTURE CAMERAS E 417.e4
BT CAMERAS
NT CINE CAMERAS
TELEVISION CAMERAS

PHOTOGRAPHY R 562.d5
RT CAMERAS E 417

REFLEX CAMERAS E 417.c3
BT STILL CAMERAS
NT SINGLE LENS REFLEX CAMERAS
TWIN LENS REFLEX CAMERAS

SINGLE LENS REFLEX CAMERAS E 417.c2
BT REFLEX CAMERAS

STEREO CAMERAS E 417.c6
BT CAMERAS

STILL CAMERAS E 417.c4
BT CAMERAS
NT INSTANT PICTURE CAMERAS
MINIATURE CAMERAS
REFLEX CAMERAS
VIEW CAMERAS

TELEVISION R 685.d5
RT television cameras E 417

TELEVISION CAMERAS E 417.e3
BT MOVING PICTURE CAMERAS
RT TELEVISION R 685

TWIN LENS REFLEX CAMERAS E 417.d2
BT REFLEX CAMERAS

UNDERWATER CAMERAS E 417.e6
BT CAMERAS
NT UNDERWATER CINE CAMERAS
RT DIVING T 473

UNDERWATER CINE CAMERAS E 417.g4
BT CINE CAMERAS
UNDERWATER CAMERAS

VIEW CAMERAS E 417.b4
SN Cameras with through-the-lens focusing and a range of movements of the lens plane relative to the film plane
UF land cameras
BT STILL CAMERAS

Figure 4 b) — Alphabetical index to arrowgraph

10 Management aspects of thesaurus construction

10.1 Methods of compilation

10.1.1 If possible, the form of the thesaurus (i.e. alphabetical listing, systematic display or graphic display) should be decided before terms are collected and considered as candidates for inclusion. Two initial approaches to thesaurus construction are possible:

a) **The deductive method**

When applying this technique, terms are extracted from documents during a preliminary stage of indexing, but no attempt is made to control the vocabulary, nor to determine relationships between terms, until a sufficient number of terms has been collected. All terms are then reviewed by a group of specialists, preferably consisting of both indexers and subject experts. These should first identify terms which represent the broadest categories, and the remaining terms should then be allocated to these categories on the basis of their logical relationships, so that the categories tend to be established on a broader-to-narrower term basis. Vocabulary control should be applied at the stage where categories are established, following the principles described in clauses 5 to 7.

b) **The inductive method**

With the inductive method, new terms are admitted into the thesaurus as soon as they have been encountered in documents. Vocabulary control is applied from the outset, and each term, as it is admitted, is designated as a member of one or more broader categories constructed on an ad hoc basis at an early stage. The thesaurus is therefore established on a narrower-to-broader term basis. Thesaurus construction is regarded from the outset as a continuous operation, and although assistance from technical experts should be sought as necessary, they need not be involved as members of a formal editorial group or committee.

10.1.2 In practice, both the deductive and inductive approaches are likely to be employed at one stage or another during the construction of a thesaurus. For example, categories of terms which were first established inductively may later be examined from a deductive viewpoint by an editorial group consisting of indexers and specialists. Both techniques are essentially empirical, and it needs to be accepted from the outset that some decisions taken during the early stages of compilation may have to be revised as further experience is gained. The addition of new terms to an indexing language frequently throws light on previous decisions concerning both the terms and their interrelationships. The compilers should check terms and hierarchies frequently to ensure that inter-term relationsips, the factoring of compound terms, etc., adhere to general principles which have been applied consistently. It is strongly recommended that expert advice should be sought when choosing between synonyms and variant forms of terms.

10.2 Recording of terms

An individual record, for example on a card, should be maintained for every term admitted into a thesaurus. This record should be made as soon as a term is admitted, and it should identify the source of the term (especially in the case of neologisms or unfamiliar terms), the names of any authorities consulted, and the date of inclusion, as well as references to synonymous, broader, narrower and related terms.

10.3 Term verification

Where possible, the correctness of a term should be verified before it is admitted into the thesaurus, and the compiler should also review relationships between each new term and other terms in the hierarchy to which it is assigned. The following types of authority should be checked when terms are accepted as candidates for inclusion:

a) standard technical dictionaries and encyclopaedias;

b) existing thesauri;

c) classification schemes.

Subject specialists, especially those with some knowledge of indexing and documentation, may also be consulted.

10.4 Specificity

The use of highly specific terms should usually be restricted to the core area of the subject field covered by the thesaurus, since an extensive use of these terms in fringe areas is likely to lead to an unwieldy and unbalanced thesaurus (see also 8.2.4). In an agency that deals with documents covering more than one domain of knowledge, it may be necessary to develop a number of specialist thesauri, each linked to, and compatible with, a general thesaurus with a lower level of specificity produced by the same agency.

10.5 Admission and deletion of terms

10.5.1 Terms and their relationships should generally reflect the ways in which language is used by specialists in the field covered by the thesaurus. If a term in common use is factored for the sake of logical consistency (see 7.3), a reference should be made from the unfactored but non-preferred form of the term.

10.5.2 When hierarchies are being established, especially during the initial stages of the inductive approach, terms which have not yet been used in indexing are frequently admitted into the thesaurus on the grounds that they can provide useful access points (for example as broader terms), and they also have potential value as preferred terms in their own right. These terms should be marked by a special symbol in the term authority file, or by a phrase such as "not yet used", and this symbol or phrase should be deleted as soon as the term is used for indexing a document.

10.5.3 Over-used terms, and terms which experience shows are used very infrequently, should be considered as candidates for deletion, since both kinds of term are generally ineffective in retrieval. In some cases an over-used term can usefully be replaced by two or more terms of greater specificity. If a term is deleted from a thesaurus but it occurs in previous indexing records, its record should be marked "for retrieval purposes only", and the date of deletion should be recorded.

10.6 The use of automatic data processing equipment

10.6.1 It has been assumed throughout this International Standard that term selection, choice of form, etc., call for intellectual decisions by indexers. Machine assistance can be employed, however, for tasks such as the following:

a) candidate terms may be identified automatically from machine-readable texts, for example titles or abstracts. The number of potential terms should first be reduced by the use of a stop list, and any remaining terms should be matched against those already recorded in the thesaurus. Unrecorded terms can be considered as potential candidates for inclusion;

b) in mechanical indexing systems, the frequency with which a term has been used in indexing (and also, perhaps, in enquiries) can be registered automatically. Terms with exceptionally high or low scores should be considered as candidates for deletion.

10.6.2 Available equipment and software may occasionally impose limitations when a thesaurus is to be reproduced from machine-held files. The following characteristics are most likely to be affected:

a) the number of characters allowed in a term;

b) the character set, including punctuation marks, available on the printer;

c) the number of levels of indention allowed in a systematic thesaurus.

Indexers should have access to equipment and programs which are capable of dealing with the special needs of a thesaurus. It should not be necessary to adapt the thesaurus to suit the facilities available on an inadequate machine.

10.7 Form and contents of a thesaurus

10.7.1 Owing to the range of displays available (see clause 9), it is not possible to stipulate a standard form of layout in a printed thesaurus. The following parts should, however, be clearly distinguished:

a) title page;

b) contents list;

c) introduction (see 10.7.2);

d) systematic or graphic displays, where appropriate;

e) alphabetical section.

10.7.2 All thesauri should contain a comprehensive introduction which clearly states

a) the purpose of the thesaurus;

b) the subject field(s) covered, with core and fringe areas separately identified;

c) the meaning of all conventions and abbreviations;

d) the total number of terms, with separate totals of preferred and non-preferred terms;

e) the rules adopted in selecting the preferred forms of indexing terms, and in deciding their interrelationships;

f) the filing rules employed, following (and quoting) an appropriate national or International Standard when possible;

g) the meanings of any punctuation marks used in non-standard ways;

h) a statement on the updating policy, and the name and address of the responsible agency to whom comments and suggestions should be sent;

i) the date when the last term was added.

Whenever possible, these points should be illustrated by examples.

10.8 Other editorial matters

10.8.1 Notification of intent

When an agency has decided to publish a new thesaurus, notification of intent should be announced in an appropriate professional journal.

10.8.2 Avoidance of duplicate work

The compiler should ascertain, through reference to one of the clearinghouses named below, whether an existing thesaurus covers the same or an overlapping domain of knowledge. Complete duplication of subject coverage is rare, but preliminary access to one or more thesauri in related fields can frequently serve as a useful starting point.

10.8.3 Deposit with a clearinghouse

A copy of the first and any subsequent edition of a thesaurus should be deposited with the appropriate national clearinghouse, and with one of the following international clearinghouses:

a) thesauri in the English language, including multilingual thesauri containing English language sections:

Thesaurus Clearinghouse
The Library
Faculty of Library Science
University of Toronto
140 St George Street
TORONTO
Ontario M5S 1A1
Canada

b) thesauri in other languages:

Instytut Informacji Naukowej,
Technicznej i Economicznej
Clearinghouse
ul. Żurawia 3/5
00-926 WARSZAWA
Poland

10.8.4 Review procedures

It is recommended that a thesaurus should be tested by means of a pilot run prior to publication. A provisional version should be distributed to a select group of users, and any suggestions for changes to terms and/or their interrelationships should be considered by the editors. The thesaurus should also be subject to a major review after a period in use, and at regular intervals thereafter, taking acount of any changes in the use of terms within the field covered by the thesaurus, and the frequency of their occurrence in indexing and/or enquiries.

Annex
Symbolization of thesaural relationships

(This annex forms part of the standard.)

A.1 The following symbols have been used throughout this International Standard to express the basic thesaural relationships:

Equivalence relationship

USE precede the preferred term

UF precedes the non-preferred term

Hierarchical relationship

TT Top term in a hierarchy

BT Broader term

Further distinctions can be achieved by the use of:

BTG Broader term (generic)

BTP Broader term (partitive)

NT Narrower or more specific term

Further distinctions can be achieved by the use of the symbols:

NTG Narrower term (generic)

NTP Narrower term (partitive)

Associative relationship

RT Related term

Equivalent symbols which represent abbreviations of terms in other languages are also encountered in published thesauri. A sample of these is given in 4.2.

A.2 These symbols have acquired, through general use, the status of "standard" conventions, but they are also language dependent to a large extent. Indexing agencies dealing with users with a range of natural languages, including those concerned with the production of multilingual thesauri, might therefore prefer a more neutral or language independent system of symbols. Those shown in the table, which have been established by ISO, are offered as a possible system:

Table — Symbols

Symbol	Significance
	Equivalence relationship
	Symbol which precedes:
→	a) the preferred term
=	b) the non-preferred term
	Hierarchical relationship
	Symbol which precedes:
<	a) the broader or superordinate term
>	b) the narrower or subordinate term
—<	c) the broader term (partitive)
>—	d) the narrower term (partitive)
	Associative relationship
–	Symbol which precedes each related term
	Conjunction
+	Symbol printed between two terms to indicate that they should be used in combination to represent a compound concept

Some of the symbols in the table do not appear on a standard typewriter keyboard, but they are available on some computer print chains, and they can be produced without difficulty in any hand-set or photocomposed thesaurus.

A.3 It is stressed that these conventions are intended to express relationships only in printed thesauri. Relationships between terms in thesauri which are held in machine-readable form (for example on disk or tape) can be expressed by any suitable combination of characters, provided that standard conventions or abbreviations are used to express these relationships in the printed output.

Bureau of Indian Standards

BIS is a statutory institution established under the *Bureau of Indian Standards Act*, 1986 to promote harmonious development of the activities of standardization, marking and quality certification of goods and attending to connected matters in the country.

Copyright

BIS has the copyright of all its publications. No part of these publications may be reproduced in any form without the prior permission in writing of BIS. This does not preclude the free use, in the course of implementing the standard, of necessary details, such as symbols and sizes, type or grade designations. Enquiries relating to copyright be addressed to the Director (Publication), BIS.

Review of Indian Standards

Amendments are issued to standards as the need arises on the basis of comments. Standards are also reviewed periodically; a standard along with amendments is reaffirmed when such review indicates that no changes are needed; if the review indicates that changes are needed, it is taken up for revision. Users of Indian Standards should ascertain that they are in possession of the latest amendments or edition by referring to the latest issue of 'BIS Handbook' and 'Standards Monthly Additions'.

This Indian Standard has been developed from Doc: No. MSD 5 (142)

Amendments Issued Since Publication

Amend No.	Date of Issue	Text Affected

BUREAU OF INDIAN STANDARDS

Headquarters:

Manak Bhavan, 9 Bahadur Shah Zafar Marg, New Delhi 110002
Telephones: 323 01 31, 323 33 75, 323 94 02

Telegrams: Manaksanstha
(Common to all offices)

Regional Offices:		Telephone
Central	: Manak Bhavan, 9 Bahadur Shah Zafar Marg NEW DELHI 110002	323 76 17, 323 38 41
Eastern	: 1/14 C.I.T. Scheme VII M, V.I.P. Road, Maniktola CALCUTTA 700054	337 84 99, 337 85 61 337 86 26, 337 91 20
Northern	: SCO 335-336, Sector 34-A, CHANDIGARH 160022	60 38 43 60 20 25
Southern	: C.I.T. Campus, IV Cross Road, CHENNAI 600113	235 02 16, 235 04 42 235 15 19, 235 23 15
Western	: Manakalaya, E9 MIDC, Marol, Andheri (East) MUMBAI 400093	832 92 95, 832 78 58 832 78 91, 832 78 92
Branches	: AHMADABAD. BANGALORE. BHOPAL. BHUBANESHWAR. COIMBATORE. FARIDABAD. GHAZIABAD. GUWAHATI. HYDERABAD. JAIPUR. KANPUR. LUCKNOW. NAGPUR. PATNA. PUNE. THIRUVANANTHAPURAM.	

Printed at Simco Printing Press, Delhi, India